

Secure Multiplex Network Coding

Ryutaroh Matsumoto

Department of Communications and Integrated Systems,
Tokyo Institute of Technology, 152-8550 Japan

Masahito Hayashi

Graduate School of Information Sciences,
Tohoku University, 980-8579 Japan
and Centre for Quantum Technologies,
National University of Singapore,
3 Science Drive 2, Singapore 117542

Abstract—In the secure network coding for multicasting, there is loss of information rate due to inclusion of random bits at the source node. We show a method to eliminate that loss of information rate by using multiple statistically independent messages to be kept secret from an eavesdropper. The proposed scheme is an adaptation of Yamamoto et al.'s secure multiplex coding [14] to the secure network coding [4], [5], [20].

Keywords—information theoretic security, network coding, secure multiplex coding, secure network coding

I. INTRODUCTION

Network coding [1] attracts much attention recently because it can offer improvements in several metrics, such as throughput and energy consumption, see [9], [10]. On the other hand, the information theoretic security [16] also attracts much attention because it offers security that does not depend on a conjectured difficulty of some computational problem.

A juncture of the network coding and the information theoretic security is the secure network coding [4], [5], which prevents an eavesdropper, called Eve, from knowing the message from the legitimate sender, called Alice, to the legitimate receivers by eavesdropping intermediate links up to a specified number in a network. It can be seen [8] as a network coding counterpart of the traditional wiretap channel coding problem considered by Wyner [21] and subsequently others [16]. In both secure network coding and coding for wiretap channels, the secrecy is realized by including random bits statistically independent of the secret message into the transmitted signal by Alice so that the secret message becomes ambiguous to Eve. The inclusion of random bits, of course, decreases the information rate. In order to get rid of the decrease in the information rate, Yamamoto et al. [14] proposed the secure multiplex coding for wiretap channels, in which there is no loss of information rate. The idea of Yamamoto et al. is as follows: Suppose that Alice has T statistically independent messages S_1, \dots, S_T . Then $S_1, \dots, S_{i-1}, S_{i+1}, \dots, S_T$ serve as the random bits making S_i ambiguous to Eve, for each i . The purpose of this paper is to do the same thing with the secure network coding as Yamamoto et al. [14].

Independently and simultaneously, Bhattad and Narayanan [3] proposed the weakly secure network coding, whose goal is also to get rid of the loss of information rate in the secure network coding. Their method [3] ensures that the mutual information between S_i and Eve's information is zero for each i . As drawbacks, the construction depends on the

network topology and coding at intermediate nodes, and the computational complexity of code construction is large. In order to remove these drawbacks, Silva and Kschischang [19] proposed the universal weakly secure network coding, in which they showed an efficient code construction that can support up to two \mathbf{F}_q -symbols in each S_i and is independent of the network topology and coding at intermediate nodes. They [19] also showed the existence of universal weakly secure network coding with more than two \mathbf{F}_q -symbols in S_i , but have not shown an explicit construction.

We shall propose a construction of secure multiplex network coding that is an adaptation of Yamamoto et al.'s idea [14] to the network coding. However, we relax an aspect of the security requirements traditionally used in the secure network coding. In previous proposals of secure network coding [4], [3], [11], [19], [20] it is required that the mutual information to the eavesdropper is exactly zero. We relax this requirement by regarding sufficiently small mutual information to be acceptable. This relaxation is similar to requiring the bit error rate to be sufficiently small instead of strictly zero. Also observe that our relaxed criterion is much stronger than one commonly used in the information theoretic security [16]. Our construction can realize arbitrary small mutual information if coding over sufficiently many time slots is allowed.

There are several reasonable models of the eavesdropper Eve. In the traditional model used in [4], [5], [11], [19], [20], Eve can arbitrarily choose the set of μ eavesdropped links after learning the structure of network coding and the set of eavesdropped links is assumed to be constant during transmission of one coding block. The secure network coding is required to leak no information with every set of μ eavesdropped links. We call this model as the traditional eavesdropping model. We shall show that the proposed scheme is universal secure in Section III-C in the sense of [19], [20] under the traditional eavesdropping model.

However, it is observed in [18] that there is difficulty in implementation over the current Internet architecture to keep the set of eavesdropped links constant even when the set of eavesdropped links is physically constant. Thus, we consider another model of the eavesdropper in which the set of eavesdropped links is statistically distributed independent of the structure of transmitter, the number of eavesdropped links is μ per unit time, and the set of eavesdropped links is allowed to be time-varying. We call this second model as

the statistical eavesdropping model. We shall also show that the mutual information is small averaged over *any probability distribution* of network coding statistically independent of any other random variables, instead of the mutual information being small with every network coding as done in [4], [5], [11], [19], [20]. Since the network coding is often constructed in the random manner [13], considering the probability distribution of network coding and requiring the averaged mutual information being small make sense. We also define a Shannon theoretic capacity region of secure multiplex network coding and show that the proposed construction can achieve that capacity region.

Although Harada and Yamamoto [11] have not explicitly stated, the adaptation of the secure multiplex coding [14] to the secure network coding [4], [5] can be done by their strongly secure network coding [11]. The difference between our proposed scheme and the previous works [3], [11], [19] is as follows:

- The computational complexity of constructing network coding is huge in [3], [11], while the computational complexity of code construction in the proposed scheme is that of selecting a random nonsingular linear matrix.
- The construction of network coding in [3], [11] depends on the underlying network topology, while the proposed scheme is independent of it and universal secure in the sense of [19], [20] under the traditional eavesdropping model (see Section III-C).
- The explicit construction of universal weakly secure network coding in [19] supports up to two \mathbf{F}_q -symbols in each secret message S_i and guarantees that the mutual information between each S_i and Eve's information is zero, while the proposed scheme has no limitation on the size of S_i and can make the mutual information between any collection $(S_i : i \in \mathcal{I})$ and Eve's information arbitrary small¹ provided that the total information rate of $(S_i : i \in \mathcal{I})$ is not too large relative to μ .
- When the total information rate of $(S_i : i \in \mathcal{I})$ is large relative to μ , the mutual information to Eve becomes positive, but [3], [19] do not evaluate how large it is nor their smallest possible value. The proposed scheme realizes asymptotically the smallest possible value of mutual information with every collection $(S_i : i \in \mathcal{I})$ *simultaneously*, as well as [11].

This paper is organized as follows: Section II reviews related results used in this paper. Section III introduces the strengthened version of the privacy amplification theorem and the proposed scheme for secure network coding, then proves the asymptotic optimality of the latter. Section IV concludes the paper.

II. PRELIMINARY

A. Model of network coding

As in [3], [4], [5], [11], [19], [20] we consider the single source multicast. The network model is either acyclic or cyclic,

¹The mutual information turned out to be exactly zero, see Appendix B.

and each link has either no delay or unit delay. We assume the linear network coding [15], and a link carries single \mathbf{F}_q symbol per time slot. The linear combination coefficients at each node are fixed so that every legitimate receiver can receive n symbols per time slot from the source. Linear combination coefficients at a node are allowed to change at each time slot except in Section III-C. We shall only consider the eavesdropper Eve and forget about the legitimate receivers. We shall propose a coding method encoding information over m time slots at the source node. Therefore, the source node transmit $(m \times n)$ \mathbf{F}_q symbols in a single coding block. Eve can eavesdrop μ links per time slot. We assume $\mu \leq n$ throughout this paper. The total number of eavesdropped links is therefore $m\mu$.

B. Two-universal hash functions

We shall use a family of two-universal hash functions [6] for the privacy amplification theorem introduced later.

Definition 1: Let \mathcal{F} be a set of functions from a finite set \mathcal{S}_1 to another finite set \mathcal{S}_2 , and F the uniform random variable on \mathcal{F} . If for any $x_1 \neq x_2 \in \mathcal{S}_1$ we have

$$\Pr[F(x_1) = F(x_2)] \leq \frac{1}{|\mathcal{S}_2|},$$

then \mathcal{F} is said to be a *family of two-universal hash functions*.

III. CONSTRUCTION OF SECURE MULTIPLEX NETWORK CODING

A. Strengthened privacy amplification theorem

In order to evaluate the mutual information to Eve when the rate of secret information is large, we need to strengthen the privacy amplification theorem originally appeared in [2], [12] as follows. The original version of the privacy amplification theorems [2], [12] cannot deduce Eq. (11) while Theorem 2 can.

Theorem 2: Let X and Z be discrete random variables on finite sets \mathcal{X} and \mathcal{Z} , respectively, and \mathcal{F} be a family of two-universal hash functions from \mathcal{X} to \mathcal{S} . Let F be the uniform random variable on \mathcal{F} statistically independent of X and Z . Then we have

$$\mathbf{E}_f \exp(\rho I(F(X); Z|F = f)) \leq 1 + |\mathcal{S}|^\rho \mathbf{E}[P_{X|Z}(X|Z)^\rho] \quad (1)$$

for all $0 \leq \rho \leq 1$, where I denotes the (conditional) mutual information as defined in [7]. We use the natural logarithm for all the logarithms in this paper, which include ones implicitly appearing in entropy and mutual information. Otherwise we have to adjust the above inequality. Proof will be given in Appendix A.

B. Description of the proposed scheme

We assume that we have T statistically independent and uniformly distributed secret messages, and that the i -th secret message is given as a random variable S_i whose realization is a column vector in $\mathbf{F}_q^{k_i}$. The sizes k_i are determined later. We shall also use a supplementary random message S_{T+1} taking values in $\mathbf{F}_q^{k_{T+1}}$ when the randomness in the encoder is insufficient to make S_i secret from Eve. We assume $mn =$

$k_1 + \dots + k_{T+1}$. Let \mathcal{L} be the set of all bijective \mathbf{F}_q -linear maps from $\prod_{i=1}^{T+1} \mathbf{F}_q^{k_i}$ to itself, and α_I be the projection from $\prod_{i=1}^{T+1} \mathbf{F}_q^{k_i}$ to $\prod_{i \in I} \mathbf{F}_q^{k_i}$ for $\emptyset \neq I \subseteq \{1, \dots, T\}$. In [17] we have shown that the family $\{\alpha_I \circ L \mid L \in \mathcal{L}\}$ is that of two-universal hash functions for all $\emptyset \neq I \subseteq \{1, \dots, T\}$. Let L be the uniform random variable on \mathcal{L} statistically independent of S_1, \dots, S_{T+1} , and arbitrary fix nonempty $I \subseteq \{1, \dots, T\}$. Define X to be the random variable $L^{-1}(S_1, \dots, S_T, S_{T+1})$. The source node sends X to its n outgoing links over m time slots. Our construction just attaches the inverse of a bijective linear function to an existing network coding.

By the assumption on Eve, her information can be expressed as BX by using a $\mu m \times mn$ matrix B over \mathbf{F}_q . We regard B as a random variable (matrix) and its probability distribution is denoted by P_B . We make no assumption on P_B except that the rank of B is at most μm and that B is independent of S_1, \dots, S_{T+1} , L and X , which means that Eve does not change B by watching the realization of L . When the random network coding is used and Eve can choose locations of up to μ eavesdropping links but cannot choose the linear coefficients of random network coding, the statistical independence assumption on B looks reasonable. From the uniformity assumption on S_i , the conditional distribution $P_{X|L}$ is uniform with every realization of L . This means that X and L are statistically independent and P_X is uniform. For fixed z , the set $\{x \in \mathbf{F}_q^{mn} \mid Bx = z\}$ has $q^{mn - \text{rank} B}$ vectors. Thus, the conditional distribution $P_{X|BX, L, B}(x|z, \ell, b)$ is the uniform distribution on the set of $q^{mn - \text{rank} b}$ elements with every triple of z , ℓ and b . We have

$$P_{X|BX, L, B}(x|z, l, b) = P_{X|BX, B}(x|z, b) = q^{-(mn - \text{rank} b)}. \quad (2)$$

Arbitrary fix nonempty $I \subseteq \{1, \dots, T\}$, denote the collection of random variables $(S_i : i \in I)$ by S_I , also fix a realization b of B , and let $k_I = \sum_{i \in I} k_i$. Under these notations, we can upper bound the mutual information $I(S_I; bX|L)$ as

$$\mathbf{E}_\ell \exp(\rho I(S_I; bX|B = b, L = \ell)) \quad (3)$$

$$\leq 1 + q^{\rho k_I} \mathbf{E}[P_{X|bX}(X|bX)^\rho] \text{ (by Theorem 2)}$$

$$= 1 + q^{\rho k_I} \mathbf{E}[(q^{-(mn - \text{rank} b)})^\rho] \text{ (by Eq. (2))}$$

$$\leq 1 + q^{\rho k_I} q^{-\rho m(n - \mu)} \text{ (because } \text{rank} b \leq \mu m \text{)}$$

$$= 1 + q^{-\rho m(n - \mu - k_I/m)} \quad (4)$$

for $0 \leq \rho \leq 1$. One can also see that

$$\mathbf{E}_\ell \rho I(S_I; bX|B = b, L = \ell) \quad (5)$$

$$\log \exp(\mathbf{E}_\ell \rho I(S_I; bX|B = b, L = \ell))$$

$$\leq \log \mathbf{E}_\ell \exp(\rho I(S_I; bX|B = b, L = \ell))$$

$$\leq \log(1 + q^{-\rho m(n - \mu - k_I/m)})$$

$$\leq q^{-\rho m(n - \mu - k_I/m)} \quad (6)$$

for $0 \leq \rho \leq 1$. Averaging Eqs. (3)–(6) over b we have

$$\mathbf{E}_{b, \ell} \rho I(S_I; bX|B = b, L = \ell) \leq q^{-\rho m(n - \mu - k_I/m)}, \quad (7)$$

$$\mathbf{E}_{b, \ell} \exp(\rho I(S_I; bX|B = b, L = \ell)) \leq 1 + q^{-\rho m(n - \mu - k_I/m)}. \quad (8)$$

Fix $C_1 > 2 \times (2^T - 1)$. Equation (7) and the Markov inequality yield that

$$\Pr \mathcal{L}_{I,1} < 1/C_1$$

for any single nonempty $I \subseteq \{1, \dots, T\}$, where $\mathcal{L}_{I,1} := \{\ell \mid \mathbf{E}_b \rho I(S_I; bX|B = b, L = \ell) > C_1 \mathbf{E}_{b, \ell} \rho I(S_I; bX|B = b, L = \ell)\}$. Thus,

$$\Pr \cup_{I: I \neq \emptyset} \mathcal{L}_{I,1} < (2^T - 1)/C_1.$$

This means that a realization ℓ of L satisfies

$$\begin{aligned} \mathbf{E}_b \rho I(S_I; bX|B = b, L = \ell) \\ \leq C_1 \mathbf{E}_{b, \ell} \rho I(S_I; bX|B = b, L = \ell) \\ \leq C_1 q^{-\rho m(n - \mu - k_I/m)} / \rho \end{aligned} \quad (9)$$

for all the $(2^T - 1)$ nonempty subsets I of $\{1, \dots, T\}$ with probability at least $1 - (2^T - 1)/C_1$. Defining another subset $\mathcal{L}_{I,2} := \{\ell \mid \mathbf{E}_b \exp(\rho I(S_I; bX|B = b, L = \ell)) > C_1 \mathbf{E}_{b, \ell} \exp(\rho I(S_I; bX|B = b, L = \ell))\}$, by Eq. (8) and the Markov inequality we obtain

$$\Pr \cup_{I: I \neq \emptyset} (\mathcal{L}_{I,1} \cup \mathcal{L}_{I,2}) < 2(2^T - 1)/C_1.$$

Therefore, a realization ℓ of L satisfies both Eq. (9) and

$$\mathbf{E}_b \exp(\rho I(S_I; bX|B = b, L = \ell)) \leq C_1 (1 + q^{-\rho m(n - \mu - k_I/m)}). \quad (10)$$

with probability at least $1 - 2 \times (2^T - 1)/C_1$.

Equation (10) implies

$$\begin{aligned} \mathbf{E}_b \frac{I(S_I; bX|B = b, L = \ell)}{m} \\ = \frac{1}{m} \mathbf{E}_b \log \exp I(S_I; bX|B = b, L = \ell) \\ \leq \frac{1}{m} \log \mathbf{E}_b \exp I(S_I; bX|B = b, L = \ell) \\ \leq \frac{\log C_1}{m\rho} + \frac{1}{m\rho} \log(1 + q^{-\rho m(n - \mu - k_I/m)}) \text{ (by Eq. (10))} \\ \leq \frac{1 + \log C_1}{m\rho} + (k_I/m - (n - \mu)) \log q, \end{aligned} \quad (11)$$

for $k_I/m - (n - \mu) \geq 0$, where in Eq. (11) we used $\log(1 + \exp(x)) \leq 1 + x$ for $x \geq 0$. Note that Eq. (9) is minimized at $\rho = 1$, because its partial derivative with respect to ρ is

$$-\frac{C_1 q^{-\rho m(n - \mu - k_I/m)}}{\rho^2} - \frac{C_1 m(n - \mu - k_I/m) \log(q) q^{-\rho m(n - \mu - k_I/m)}}{\rho},$$

which is negative for all $0 < \rho \leq 1$. By the same reason Eq. (11) is also minimized at $\rho = 1$.

Fix C_2 with $C_2 > 2 \times (2^T - 1)$. By the same argument as above, we see that for a realization ℓ of L satisfying both Eqs. (9) and (11), with probability at least

$$1 - 2 \times (2^T - 1)/C_2, \quad (12)$$

a realization b of B makes

$$I(S_I; bX|B = b, L = \ell) \leq C_1 C_2 q^{-\rho m(n - \mu - k_I/m)} / \rho, \quad (13)$$

$$\frac{I(S_I; BX|B=b, L=\ell)}{m} \leq \frac{1 + \log C_2 + \log C_1}{m\rho} + (k_I/m - (n-\mu)) \log q. \quad (14)$$

Even when we choose large C_2 and C_1 and make the probability of b and ℓ satisfying Eqs. (13) and (14) large, we can make the upper bound (13) arbitrary small by increasing m . This observation enables us to use a fixed bijective linear function ℓ that is agreed between the legitimate sender and the receivers in advance. The legitimate receiver can apply the agreed ℓ to the received information in order to restore the original messages S_i .

One can easily see that if $\rho = 1$ and $k_I \leq m(n-\mu-\delta_I)$ with $\delta_I > 0$, then the upper bound (13) on the mutual information exponentially converges to zero as $m \rightarrow \infty$. This means that we can transmit $(n-\mu) \mathbf{F}_q$ -symbols in the i -th secret message per time slot with arbitrary small eavesdropped information when the encoding is allowed to be done over sufficiently many time slots. On the other hand, if $m(n-\mu) \leq k_I \leq m(R_I - \delta_I)$ with $\delta_I > 0$ and $R_I > 0$, then by the upper bound (14) we see the mutual information per symbol is upper bounded by

$$\frac{I(S_I; BX|B=b, L=\ell)}{m} < R_I - (n-\mu)$$

for sufficiently large m . In Section III-D we shall show that the proposed scheme is asymptotically optimal by capacity consideration.

Remark 3: The meanings of C_1 and C_2 are as follows: At Eqs. (7) and (8), there might not exist a realization ℓ of L that satisfies Eqs. (7) and (8) for all subsets \mathcal{I} of $\{1, \dots, T\}$ simultaneously. By sacrificing the tightness of the upper bounds, we ensure the existence of ℓ satisfying Eqs. (9) and (10). At Eqs. (9) and (10), realizations b of B might satisfy Eqs. (9) and (10) with unacceptably low probability. Again by sacrificing the tightness of the upper bounds, we ensure that the eavesdropping matrix B satisfies Eqs. (13) and (14) with comfortably high probability.

C. Security analysis of the proposed scheme under the traditional eavesdropping model

In the preceding study of secure network coding [4], [5], [11], [19], [20], it is assumed that

- the eavesdropper Eve can choose μ eavesdropped links per unit time after learning the structure of network coding, and
- the set of eavesdropped links is constant during transmission of one coding block.

In Section I we called the above assumption as the traditional eavesdropping model. Under the above assumption, the number of possible sets of eavesdropped links is constant, say C_E , independent of m . If we take the random variable B according to the uniform distribution on the sets of eavesdropped links and the probability of some event of B is larger than $1 - 1/C_E$ then that probability must be one. Set C_2 in Eq. (12) such that $1 - 2 \times (2^T - 1)/C_2 > 1 - 1/C_E$, then we can see that Eqs. (13) and (14) hold with every realization of B . In addition to this, the proposed construction of coding does not depend

on the network topology nor coding at intermediate nodes. In that sense, the proposed scheme is universal secure [19], [20] except that the proposed scheme can make the mutual information arbitrary small², while [19], [20] make the mutual information exactly zero.

D. Capacity consideration of the proposed scheme

Firstly, let us define the achievable rate tuple and the capacity region as in [7].

Definition 4: Suppose that we are given a sequence of $\mu m \times mn$ random matrices B_m whose distribution P_{B_m} has no restriction except that the rank of B_m is always μm and that B_m is independent of any other random variables. Let e_m be a stochastic encoder from $\prod_{i=1}^T \mathbf{F}_q^{mk_{i,m}}$ to \mathbf{F}_q^{mn} , d_m either stochastic or deterministic decoder from \mathbf{F}_q^{mn} to $\prod_{i=1}^T \mathbf{F}_q^{mk_{i,m}}$, and $S_{i,m}$ the uniform random variable on $\mathbf{F}_q^{mk_{i,m}}$, for $i = 1, \dots, T$ and $m = 1, 2, \dots$. If

$$\lim_{m \rightarrow \infty} \Pr[(S_{1,m}, \dots, S_{T,m}) \neq d_m(e_m(S_{1,m}, \dots, S_{T,m}))] = 0,$$

$$\limsup_{m \rightarrow \infty} I(S_{\mathcal{I},m}; B_m e_m(S_{1,m}, \dots, S_{T,m}) | B_m) / m$$

$$\leq \max \left\{ 0, -(n-\mu) + \sum_{i \in \mathcal{I}} R_i \right\},$$

$$\liminf_{m \rightarrow \infty} \kappa_{i,m} \geq R_i,$$

for all nonempty subsets $\mathcal{I} \subseteq \{1, \dots, T\}$, then the rate tuple (R_1, \dots, R_T) is said to be *achievable* for the secure multiplex network coding with T secret messages and up to μ eavesdropped links per time slots represented by random matrices B_m , where $S_{\mathcal{I},m}$ denotes the collection $(S_{i,m} : i \in \mathcal{I})$ of secret messages $S_{i,m}$. We also define the *capacity region* of \mathbf{F}_q -linear secure multiplex network coding as the closure of such rate tuples (R_1, \dots, R_T) over all the sequences of encoders and decoders.

Theorem 5: The capacity region \mathbf{F}_q -linear secure multiplex network coding is given by rate tuples (R_1, \dots, R_T) such that

$$0 \leq R_i,$$

$$\sum_{i=1}^T R_i \leq n.$$

Proof. The fact that every rate tuple given by the above equation is achievable is already proved in the previous section under stronger requirements, namely $\Pr[(S_{1,m}, \dots, S_{T,m}) \neq d_m(e_m(S_{1,m}, \dots, S_{T,m}))] = 0$ for all m , $\limsup_{m \rightarrow \infty} I(S_{\mathcal{I},m}; B_m e_m(S_{1,m}, \dots, S_{T,m}) | B_m) / m \leq \max\{0, \sum_{i \in \mathcal{I}} R_i - (n-\mu)\}$, and $\lim_{m \rightarrow \infty} I(S_{\mathcal{I},m}; B_m e_m(S_{1,m}, \dots, S_{T,m}) | B_m) = 0$ for \mathcal{I} with $\sum_{i \in \mathcal{I}} R_i < (n-\mu)$. We have to show the so-called converse part of the coding theorem. Fix an arbitrary nonempty subset $\mathcal{I} \subseteq \{1, \dots, T\}$ and a sequence of random matrices B_m , and suppose that we have a sequence of stochastic encoders as defined in Definition 4 with

$$\liminf_{m \rightarrow \infty} \sum_{i \in \mathcal{I}} \kappa_{i,m} \geq \sum_{i \in \mathcal{I}} R_i + \delta \quad (15)$$

²The mutual information turned out to be exactly zero and our scheme is exactly universal secure in the sense of [19], [20], see Appendix B.

for $\delta > 0$ and suppose also that

$$\limsup_{m \rightarrow \infty} I(S_{I,m}; B_m e_m(S_{1,m}, \dots, S_{T,m}) | B_m) / m \leq \max \left\{ 0, \sum_{i \in \mathcal{I}} R_i - (n - \mu) \right\}. \quad (16)$$

By Eq. (16) there exists a sequence of $\mu m \times mn$ matrices b_m such that

$$\begin{aligned} & \limsup_{m \rightarrow \infty} I(S_{I,m}; b_m e_m(S_{1,m}, \dots, S_{T,m})) / m \\ & \leq \limsup_{m \rightarrow \infty} I(S_{I,m}; B_m e_m(S_{1,m}, \dots, S_{T,m}) | B_m) / m \\ & \leq \max \left\{ 0, \sum_{i \in \mathcal{I}} R_i - (n - \mu) \right\}. \end{aligned} \quad (17)$$

For every m , define a_m to be an $\mu m \times \mu m$ matrix and c_m to be an $mn \times mn$ matrix such that $a_m b_m c_m$ is a matrix of the horizontal concatenation of the $\mu m \times \mu m$ identity matrix and the zero matrix. By Eq. (17) and the data processing inequality we have

$$\begin{aligned} & \limsup_{m \rightarrow \infty} I(S_{I,m}; a_m b_m c_m e_m(S_{1,m}, \dots, S_{T,m})) / m \\ & \leq \max \left\{ 0, \sum_{i \in \mathcal{I}} R_i - (n - \mu) \right\}. \end{aligned} \quad (18)$$

Define $X_m^{(1)}$ as the first μm components in the random vector $e_m(S_{1,m}, \dots, S_{T,m})$, and $X_m^{(2)}$ as the remaining components in $e_m(S_{1,m}, \dots, S_{T,m})$. We have

$$\begin{aligned} & H(S_{I,m} | e_m(S_{1,m}, \dots, S_{T,m})) \\ & = H(S_{I,m} | X_m^{(1)}, X_m^{(2)}) \\ & = H(S_{I,m}) - I(S_{I,m}; X_m^{(1)}, X_m^{(2)}) \\ & = H(S_{I,m}) - I(S_{I,m}; X_m^{(1)}) - I(S_{I,m}; X_m^{(2)} | X_m^{(1)}) \text{ (by the chain rule)} \\ & = H(S_{I,m}) - I(S_{I,m}; a_m b_m c_m e_m(S_{1,m}, \dots, S_{T,m})) \\ & \quad - I(S_{I,m}; X_m^{(2)} | X_m^{(1)}) \\ & \geq H(S_{I,m}) - I(S_{I,m}; b_m e_m(S_{1,m}, \dots, S_{T,m})) - I(S_{I,m}; X_m^{(2)} | X_m^{(1)}) \\ & \geq H(S_{I,m}) - I(S_{I,m}; b_m e_m(S_{1,m}, \dots, S_{T,m})) - H(X_m^{(2)}) \\ & \geq m \sum_{i \in \mathcal{I}} \kappa_{i,m} \log q - I(S_{I,m}; b_m e_m(S_{1,m}, \dots, S_{T,m})) \\ & \quad - m(n - \mu) \log q \\ & \geq m \left[\left(\sum_{i \in \mathcal{I}} \kappa_{i,m} - n + \mu \right) \log q - I(S_{I,m}; b_m e_m(S_{1,m}, \dots, S_{T,m})) / m \right] \\ & \geq m \left[\left(\sum_{i \in \mathcal{I}} \kappa_{i,m} - n + \mu \right) \log q \right. \\ & \quad \left. - I(S_{I,m}; B_m e_m(S_{1,m}, \dots, S_{T,m}) | B_m) / m \right], \end{aligned}$$

where H denotes the (conditional) entropy as defined in [7]. By abuse of notation, re-define α_I to be the projection from $\prod_{i=1}^T \mathbf{F}_q^{k_i}$ to $\prod_{i \in \mathcal{I}} \mathbf{F}_q^{k_i}$ for $\emptyset \neq \mathcal{I} \subseteq \{1, \dots, T\}$. By using Fano's inequality [7, Theorem 2.10.1]

$$\begin{aligned} & \Pr[(S_{1,m}, \dots, S_{T,m}) \neq d_m(e_m(S_{1,m}, \dots, S_{T,m}))] \\ & \geq \Pr[S_{I,m} \neq \alpha_I(d_m(e_m(S_{1,m}, \dots, S_{T,m})))] \end{aligned}$$

$$\begin{aligned} & \geq \frac{H(S_{I,m} | e_m(S_{1,m}, \dots, S_{T,m})) - 1}{\log |\prod_{i \in \mathcal{I}} \mathbf{F}_q^{m \kappa_{i,m}}|} \\ & \geq \frac{1}{\sum_{i \in \mathcal{I}} \kappa_{i,m} \log q} \left[\left(\sum_{i \in \mathcal{I}} \kappa_{i,m} - n + \mu \right) \log q \right. \\ & \quad \left. - \frac{1}{m} - \frac{I(S_{I,m}; B_m e_m(S_{1,m}, \dots, S_{T,m}) | B_m)}{m} \right]. \end{aligned} \quad (19)$$

By Eqs. (15), (16) and (19) we can see that $\limsup_{m \rightarrow \infty} \Pr[(S_{1,m}, \dots, S_{T,m}) \neq d_m(e_m(S_{1,m}, \dots, S_{T,m}))] \geq \delta / (\delta + \sum_{i \in \mathcal{I}} R_i) > 0$. This shows that the limit of mutual information $I(S_{I,m}; B_m e_m(S_{1,m}, \dots, S_{T,m}) | B_m) / n$ cannot be lower than $\sum_{i \in \mathcal{I}} R_i - (n - \mu)$ while keeping the sum of information rates $\sum_{i \in \mathcal{I}} \kappa_i$ strictly larger than $\sum_{i \in \mathcal{I}} R_i$. ■

IV. CONCLUSION

In the secure network coding, there was loss of information rate due to inclusion of random bits at the source node. In this paper, we have shown that a method to eliminate that loss of information rate by using multiple statistically independent messages to be kept secret from an eavesdropper, and called the proposed scheme *secure multiplex network coding*. The proposed scheme is an adaptation of Yamamoto et al.'s secure multiplex coding [14] to the secure network coding [4], [5], [20].

ACKNOWLEDGMENT

The authors thank anonymous reviewers of NetCod 2011 for carefully reading the initial manuscript and pointing out its shortcomings. The first author would like to thank Prof. Hiro-suke Yamamoto to teach him the secure multiplex coding, Dr. Shun Watanabe to point out the relation between the proposed scheme and [11], Mr. Jun Kurihara to point out the relation between the proposed scheme and [19], Dr. Jun Muramatsu and Prof. Tomohiro Ogawa for the helpful discussion on the universal coding. A part of this research was done during the first author's stay at the Institute of Network Coding, the Chinese University of Hong Kong, and he greatly appreciates the hospitality by Prof. Raymond Yeung. This research was partially supported by the MEXT Grant-in-Aid for Young Scientists (A) No. 20686026 and (B) No. 22760267, and Grant-in-Aid for Scientific Research (A) No. 23246071. The Center for Quantum Technologies is funded by the Singapore Ministry of Education and the National Research Foundation as part of the Research Centres of Excellence programme.

APPENDIX A

PROOF OF THEOREM 2

In order to show Theorem 2, we introduce the following lemma.

Lemma 6: Under the same assumption as Theorem 2, we have

$$\mathbf{E}_f \exp(-\rho H(F(X); Z | F = f)) \leq |S|^{-\rho} + \mathbf{E}[P_{X|Z}(X|Z)^\rho] \quad (20)$$

for $0 \leq \rho \leq 1$.

Proof of Theorem 2.

$$\mathbf{E}_f \exp(\rho I(F(X); Z | F = f))$$

$$\begin{aligned}
&= \mathbf{E}_f \exp(\rho H(F(X)|F=f) - \rho H(F(X); Z|F=f)) \\
&\leq \mathbf{E}_f |S|^\rho \exp(-\rho H(F(X); Z|F=f)) \\
&\leq \mathbf{E}_f |S|^\rho (|S|^{-\rho} + \mathbf{E}[P_{X|Z}(X|Z)^\rho]) \text{ (by Eq. (20))} \\
&= 1 + |S|^\rho \mathbf{E}[P_{X|Z}(X|Z)^\rho]
\end{aligned}$$

Proof of Lemma 6. Fix $z \in \mathcal{Z}$. The concavity of x^ρ for $0 \leq \rho \leq 1$ implies

$$\begin{aligned}
&\mathbf{E}_f \sum_{s \in \mathcal{S}} P_{f(X)|Z}(s|z)^{1+\rho} \\
&= \sum_{x \in \mathcal{X}} P_{X|Z}(x|z) \mathbf{E}_f \left(\sum_{x' \in f^{-1}(x)} P_{X|Z}(x'|z) \right)^\rho \\
&\leq \sum_{x \in \mathcal{X}} P_{X|Z}(x|z) \underbrace{\left(\mathbf{E}_f \sum_{x' \in f^{-1}(x)} P_{X|Z}(x'|z) \right)^\rho}_{(*)}. \quad (21)
\end{aligned}$$

Since f is chosen from a family of two-universal hash functions defined in Definition 1, we have

$$\begin{aligned}
(*) &\leq P_{X|Z}(x|z) + \sum_{x \neq x' \in \mathcal{X}} \frac{P_{X|Z}(x'|z)}{|S|} \\
&\leq P_{X|Z}(x|z) + |S|^{-1}.
\end{aligned}$$

Since any two positive numbers x and y satisfy $(x+y)^\rho \leq x^\rho + y^\rho$ for $0 \leq \rho \leq 1$, we have

$$(P_{X|Z}(x|z) + |S|^{-1})^\rho \leq P_{X|Z}(x|z)^\rho + |S|^{-\rho}. \quad (22)$$

By Eqs. (21) and (22) we can see

$$\mathbf{E}_f \sum_{s \in \mathcal{S}} P_{f(X)|Z}(s|z)^{1+\rho} \leq \sum_{x \in \mathcal{X}} P_{X|Z}(x|z)^{1+\rho} + |S|^{-\rho}.$$

Taking the average over Z of the both sides of the last equation, we have

$$\mathbf{E}_f \mathbf{E}_{XZ} P_{f(X)|Z}(f(X)|Z)^\rho \leq \mathbf{E}_{XZ} P_{X|Z}(X|Z)^\rho + |S|^{-\rho}. \quad (23)$$

Define $g(\rho) = \mathbf{E}_{XZ} P_{f(X)|Z}(f(X)|Z)^\rho$ as a function of ρ with fixed f and P_{XZ} , and $h(\rho) = \log g(\rho)$. We have

$$\begin{aligned}
g'(\rho) &= \mathbf{E}_{XZ} P_{f(X)|Z}(f(X)|Z)^\rho \log P_{f(X)|Z}(f(X)|Z), \\
g''(\rho) &= \mathbf{E}_{XZ} P_{f(X)|Z}(f(X)|Z)^\rho (\log P_{f(X)|Z}(f(X)|Z))^2, \\
h'(\rho) &= g'(\rho)/g(\rho), \\
h''(\rho) &= \frac{g''(\rho)g(\rho) - [g'(\rho)]^2}{g(\rho)^2}.
\end{aligned}$$

Define (X', Z') to be the random variables that have the same joint distribution as (X, Z) and statistically independent of X and Z . To examine the sign of $h''(\rho)$ we compute

$$\begin{aligned}
&g''(\rho)g(\rho) - [g'(\rho)]^2 \\
&= \mathbf{E}_{XZX'Z'} P_{f(X)|Z}(f(X), Z)^\rho P_{f(X)|Z}(f(X'), Z')^\rho \\
&\quad [(\log P_{f(X)|Z}(f(X)|Z))^2 - \log P_{f(X)|Z}(f(X)|Z) \log P_{f(X)|Z}(f(X')|Z')] \\
&= \frac{1}{2} \mathbf{E}_{XZX'Z'} P_{f(X)|Z}(f(X), Z)^\rho P_{f(X)|Z}(f(X'), Z')^\rho \\
&\quad [(\log P_{f(X)|Z}(f(X)|Z))^2 + (\log P_{f(X)|Z}(f(X')|Z'))^2 \\
&\quad - 2 \log P_{f(X)|Z}(f(X)|Z) \log P_{f(X)|Z}(f(X')|Z')]
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} \mathbf{E}_{XZX'Z'} P_{f(X)|Z}(f(X), Z)^\rho P_{f(X)|Z}(f(X'), Z')^\rho \\
&\quad [\log P_{f(X)|Z}(f(X)|Z) - \log P_{f(X)|Z}(f(X')|Z')]^2 \\
&\geq 0.
\end{aligned}$$

This means that $h''(\rho) \geq 0$ and $h(\rho)$ is convex. We can see

$$\begin{aligned}
\mathbf{E}_{XZ} P_{f(X)|Z}(f(X)|Z)^\rho &= \exp(h(\rho)) \\
&\geq \exp(\underbrace{h(0)}_{=0} + \rho h'(0)) \\
&= \exp(-\rho H(f(X)|Z)). \quad (24)
\end{aligned}$$

By Eqs. (23) and (24) we see that Eq. (20) holds. ■

REFERENCES

- [1] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inform. Theory*, vol. 46, no. 4, pp. 1204–1206, Jul. 2000.
- [2] C. H. Bennett, G. Brassard, C. Crépeau, and U. M. Maurer, "Generalized privacy amplification," *IEEE Trans. Inform. Theory*, vol. 41, no. 6, pp. 1915–1923, Nov. 1995.
- [3] K. Bhattad and K. R. Narayanan, "Weakly secure network coding," in *Proc. NetCod 2005*, Riva del Garda, Italy, Apr. 2005.
- [4] N. Cai and R. W. Yeung, "Secure network coding," in *Proc. ISIT 2002*, Lausanne, Switzerland, Jul. 2002, p. 323.
- [5] —, "Secure network coding on a wiretap network," *IEEE Trans. Inform. Theory*, vol. 57, no. 1, pp. 424–435, Jan. 2011.
- [6] J. L. Carter and M. N. Wegman, "Universal classes of hash functions," *J. Comput. System Sci.*, vol. 18, no. 2, pp. 143–154, Apr. 1979.
- [7] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Wiley Interscience, 2006.
- [8] S. Y. El Rouayheb and E. Soljanin, "On wiretap networks II," in *Proc. ISIT 2007*, Nice, France, Jun. 2007, pp. 551–555.
- [9] C. Fragouli and E. Soljanin, *Network Coding Applications*. NOW Publishers, 2007.
- [10] —, *Network Coding Fundamentals*. NOW Publishers, 2007.
- [11] K. Harada and H. Yamamoto, "Strongly secure linear network coding," *IEICE Trans. Fundamentals*, vol. E91-A, no. 10, pp. 2720–2728, Oct. 2008.
- [12] M. Hayashi, "Exponential decreasing rate of leaked information in universal random privacy amplification," *IEEE Trans. Inform. Theory*, vol. 57, no. 6, pp. 3989–4001, Jun. 2011.
- [13] T. Ho, M. Medard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and B. Leong, "A random linear network coding approach to multicast," *IEEE Trans. Inform. Theory*, vol. 52, no. 10, pp. 4413–4430, Oct. 2006.
- [14] D. Kobayashi, H. Yamamoto, and T. Ogawa, "How to attain the ordinary channel capacity securely in wiretap channels," in *Proc. 2005 IEEE Information Theory Workshop on Theory and Practice in Information-Theoretic Security*, Oct. 2005, pp. 13–18, arXiv:cs/0509047.
- [15] S.-Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Trans. Inform. Theory*, vol. 49, no. 2, pp. 371–381, Feb. 2003.
- [16] Y. Liang, H. V. Poor, and S. Shamai (Shitz), *Information Theoretic Security*. Hanover, MA, USA: NOW Publishers, 2009.
- [17] R. Matsumoto and M. Hayashi, "Secure multiplex coding with a common message," in *Proc. ISIT 2011*, Saint-Petersburg, Russia, Jul. 2011, to appear, arXiv:1101.4036.
- [18] E. Shioji, R. Matsumoto, and T. Uyematsu, "Vulnerability of MRD-code-based universal secure network coding against stronger eavesdroppers," *IEICE Trans. Fundamentals*, vol. E93-A, no. 11, pp. 2026–2033, Nov. 2010.
- [19] D. Silva and F. R. Kschischang, "Universal weakly secure network coding," in *Proc. ITW 2009*, Volos, Greece, Jun. 2009, pp. 281–285.
- [20] —, "Universal secure network coding via rank-metric codes," *IEEE Trans. Inform. Theory*, vol. 57, no. 2, pp. 1124–1135, Feb. 2011.
- [21] A. D. Wyner, "The wire-tap channel," *Bell System Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

This appendix is a note after submission of the final version to Proc. NetCod 2011. In Section III-C we claimed that the mutual information to Eve can be arbitrary small. This means that the mutual information can be made exactly zero for every eavesdropping matrix b . The reason is as follows: For fixed b and ℓ , we have

$$I(S_I; BX|B = b, L = \ell) = H(S_I|B = b, L = \ell) - H(S_I|bX, L = \ell). \quad (25)$$

The first term $H(S_I|B = b, L = \ell)$ is an integer multiple of $\log q$ since S_I is assumed to have the uniform distribution. For fixed b and ℓ , we have $bX = b\ell^{-1}(S_1, \dots, S_{T+1})$. For a given realization bx of bX , the set of solutions s such that $bx = b\ell^{-1}s$ is written as $\ker(b\ell^{-1}) + \text{some vector } v$. This means that the set of possible candidates of S_I given realization bx of bX is written as $\alpha_I(\ker(b\ell^{-1})) + \alpha_I(v)$, and S_I given realization bx is uniformly distributed on $\alpha_I(\ker(b\ell^{-1})) + \alpha_I(v)$. Since the cardinality of $\alpha_I(\ker(b\ell^{-1})) + \alpha_I(v)$ is independent of X for fixed b and ℓ , the second term $H(S_I|bX, L = \ell)$ is also an integer multiple of $\log q$. Therefore, if Eq. (13) holds for every B as verified in Section III-C and the RHS of Eq. (13) is $< \log q$, then the LHS of Eq. (13) must be zero.

In Section I we overlooked the relevant research result by Cai (“Valuable messages and random outputs of channels in linear network coding,” Proc. ISIT 2009, Seoul, Korea, Jun. 2009, pp. 413–417). Cai proved that random linear network coding gives the strongly secure network coding in the sense of [11] with arbitrarily high probability with sufficiently large finite fields. The advantages of the present result over Cai’s result are (1) we do not have to change encoding at intermediate nodes, (2) our construction is universal secure in the sense of [19], [20], and (3) much smaller finite fields can be used than Cai’s result.